Are VNTR co-localizing with breast cancer-associated SNPs?

Martin Leclerc[1], Susan L. Neuhausen[2], Hagit Schayek[3], Yael Laitman[3], Antoniou C. Antonis[1], and Eitan Friedman[3,4]

From the [1]Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK; the [2]Department of Population Sciences, Beckman Research Institute, City of Hope, Duarte, CA, USA; and the [3]OncogenetIcs unit at the Sheba Medical Center, and the [4]Sackler school of Medicine, Tel-Aviv University, Tel-Aviv Israel.

Address correspondence to

Eitan Friedman MD PhD
Director, Oncogenetics Unit, Institute of Human Genetics
Sheba Medical Center, Tel-Hashomer, 52621 Israel
Tel +972-3-530-3173 FAX +972-3-535-7308
E mails – feitan@post.tau.ac.il or eitan.friedman@sheba.health.gov.il

Abstract

Purpose: Several common genetic variants (single nucleotide polymorphisms, SNPs) have been shown to be associated with breast cancer (BC) risk in the general population, and to modify BC risk for *BRCA1* and *BRCA2* mutation carriers. Co-localization of variable number of tandem repeats (VNTR) with these BC-associated SNPS has not been comprehensively studied.

Methods: Cross referencing of genome-wide VNTR with the known BC genome-wide association studies (GWAS) SNPs significantly associated with increased risk for developing breast cancer was carried out. Analysis was based on the overlap between the VNTR and 10-kb windows around these BC-susceptibility SNPs.

Results: Cross referencing of the 1.2 million TR with the 161 known BC- associated SNPs in the general population led to 690 matches. Of those, in 17 VNTRs, the SNP was within the VNTR. Analysis restricted to loci known to modify BC penetrance in *BRCA1* (n=31) and *BRCA2* (n=33) mutation carriers led to 139 and 170 co-localization matches, respectively. For these, none of the SNPs were within the VNTR. The distances between the SNPs and the VNTRs were not significantly different from what was expected to occur by chance alone (p=0.61; p=0.44; p=0.25, respectively).

Conclusion: There is no evidence that VNTRs co-localize with currently reported SNP tagged BC GWAS loci.

Key words: Tandem repeats; Breast cancer susceptibility; GWAS based SNPs; Penetrance; *BRCA1 BRCA2*

Introduction

Several genetic variants that contribute to breast cancer (BC) susceptibility have been identified, with *BRCA1* and *BRCA2* being the predominant, clinically relevant genes [1]. In some families exhibiting familial aggregation of BC, BC is diagnosed at earlier age with each successive generation [2]. This phenomenon has largely been attributed to a birth cohort effect [3]. However, it is plausible that these observed successive generational younger ages at diagnosis may reflect a genetic-based phenomenon called anticipation. Anticipation, predominantly noted in neurodegenerative disorders, implies that there are dynamic mutations particularly affecting trinucleotide repeats expansion that impact the phenotype and age at diagnosis [4]. Previous studies reported that in some cases, anticipation can clinically be defined in subset of familial BC cases [2], yet not all studies concur [5]. Furthermore, a similar, anticipation-compatible phenomenon has been reported in some families carrying mutant *BRCA1/2* alleles [2, 6]. While *BRCA1/2* germline mutation carriers are at significantly increased risk for developing breast and/or ovarian cancer, penetrance is incomplete [7]. Combined with the variable age at diagnosis of identical mutation carriers, incomplete penetrance is suggestive of the existence of modifier factors-genetic and environmental [8]. Several single nucleotide polymorphisms (SNPs) have been shown to be associated with breast or ovarian cancer risks in *BRCA1* and *BRCA2* mutation carriers (http://apps.ccge.medschl.cam.ac.uk/consortia/cimba//). Although the risk associated with individual variants is low, in combination they can result to large differences in absolute risk for mutation carriers with implications for risk management [9].

In the human genome, repetitive sequences comprise approximately half of the genome. Tandem repeats are subsets of these repetitive sequences and while the majority of tandem repeats are monomorphic, some are polymorphic, known as variable number of tandem repeats (VNTRs). The association of VNTRs and BC risk in the average risk population and as modifiers of *BRCA1/2* gene mutations has previously been reported in the context of "candidate genes" [e.g., 10]. However, studies at a genome-wide scale are relatively scarce. In this study, we are investigating the co-localization of VNTRs with significantly-associated signals for BC risk.

Methods

The Tandem Repeat database (TRDB- https://tandem.bu.edu/cgi-bin/trdb/trdb.exe?taskid=0) was accessed. This database contains nearly 1.2 million tandem repeats. SNPs that have reached a significance level of $p < 5e\text{-}08$ were considered for the general population [11] SNPs that were previously reported to be associated with BC risk for *BRCA1/2* mutation carriers ("modifiers") [12] were separately analyzed. A SNP was considered as a risk modifier in *BRCA1/2x* mutation carriers if (1): statistical significance was reached for the general population ($p < 5e\text{-}08$) as was previously described [9] and the *BRCA1/2* carriers ($p < 0.05$), and (2): the effects estimated for these two populations were in the same direction. The resulting three SNP sets (general population, *BRCA1* carriers, *and BRCA2* carriers) were cross- referenced with the TRDB. The sum of the distances between SNPs and VNTRs, denoted as D_0, was calculated for the 10kb-windows around SNPs. A resampling method similar to the local annotation shifting procedure [13] was used to assess whether there were significantly more VNTRs in the vicinity of the known BC susceptibility SNPs, by approximating the null distribution of the distances. Briefly, this involves repeating the following algorithm 10,000 times: (I) shifting all the VNTRs by δ base pairs from their initial position, where δ is a random integer between 1 and 10,000; (II) computing D_b, the sum of the distances for the shifted pattern of co-localization. The p-value for the test of co-localization was then set equal to the proportions of D_b's exceeding D_0. Mapping of the resulting SNPs and tandem repeats to the nearest genes was extracted from Michailidou et al. 2017[11] and Milne et al 2017 [12].

Results

General population- Overall, 161 SNPs that showed a significant association with BC risk for the general population were used for cross referencing. One hundred and forty six SNPs showed matching with at least one VNTR, with a total of 676 unique matching VNTRs (see Table 1 and supplementary table 1A and 1B). However, the pattern of co-localization was not significant (p=0.61). In 17 of the matches, the SNP was within the VNTR and in 5 matches, the SNP was within a tandem repeat of the Short Tandem Repeats in Regulatory Regions

Table (STaRRRT) (14). The size of the "unique" 676 TRs identified varied between 25 and 2713 base pairs, with a median of 46.5 base-pairs.

BRCA1 mutation carriers- Overall, in the CIMBA data base and using the oncoarray data, 31 SNPs that showed a significant association with BC risk for *BRCA1* mutation carriers were included. Twenty-eight SNPs showed matching with at least one VNTR, with a total of 139 unique matching VNTRs (Table 1 and supplementary table 2A and 2B). However, the pattern of co-localization was not significant (p=0.44). No SNPs were within the VNTR. The size of the unique 139 VNTRs identified varied between 25 and 2713 base-pairs with a median of 55 base-pairs.

BRCA2 mutation carriers- Overall, 33 SNPs that showed a significant association with BC risk for *BRCA2* mutation carriers, derived from the same sources as for *BRCA1*, were included. Thirty-two SNPs showed matching with at least one VNTR with a total of 158 unique matching VNTRs (Table 1 and supplementary table 23 and 3B). However, the pattern of co-localization was not significant (p=0.25). No SNPs were within the VNTR. The size of the unique 158 VNTRs identified varies between 25 and 2713 base-pairs with a median of 48 base-pairs.

Discussion

The present study investigated co-localization of VNTRs with the most significantly associated SNP in each region that exhibited genome-wide significance levels, and no statistically significant co-localization was found between VNTRs and SNPs associated with BC risk or SNPs associated with BC risk modification for *BRCA1* or *BRCA2* mutation carriers. These bioinformatics findings reduce the likelihood that VNTRs are involved in BC susceptibility or in modifying *BRCA* penetrance that can be tagged by SNPs identified through GWAS. However, because these SNPs do not necessarily represent the causal or the key "functional" variants at each locus, a more informative approach would be to investigate the set of all candidate causal SNPs in each region ("credible" set of variants); these data are not currently available for all the known BC susceptibility and risk loci. It is also possible that VNTRs play a role in BC predisposition but a GWAS design cannot identify VNTRs of

risk; not all VNTRs will be "tagged" by SNPs. Secondly, they may play a role through epigenetic changes that primarily affect VNTR, similar to what has been reported in neurodegenerative diseases with "dynamic mutations" [15]. These results do not rule out such effects. Further investigation to elucidate the role if any that VNTR play in BC susceptibility is needed.

Conflict of Interest: All authors declare that they have no conflict of interest

Acknowledgement: SLN is partially supported by the Morris and Horowitz Families Endowed Professorship. This project was performed with the help of the CIMBA consortium and its contributing members.

References

1. Kobayashi H, Ohno S, Sasaki Y, Matsuura M (2013) Hereditary breast and ovarian cancer susceptibility genes. Oncol Rep 30(3):1019-1029. doi: 10.3892/or.2013.2541
2. Litton JK1, Ready K, Chen H, Gutierrez-Barrera A, Etzel CJ, Meric-Bernstam F, Gonzalez-Angulo AM, Le-Petross H, Lu K, Hortobagyi GN, Arun BK (2012) Earlier age of onset of BRCA mutation-related cancers in subsequent generations. Cancer 118(2):321-325. doi: 10.1002/cncr.26284.
3. Gabai-Kapara E, Lahad A, Kaufman B, Friedman E, Segev S, Renbaum P, Beeri R, Gal M, Grinshpun-Cohen J, Djemal K, Mandell JB, Lee MK, Beller U, Catane R, King MC, Levy-Lahad E (2014) Population-based screening for breast and ovarian cancer risk due to BRCA1 and BRCA2.Proc Natl Acad Sci U S A 111(39):14205-14210. doi: 10.1073/pnas.1415979111.
4. Gövert F, Schneider SA (2013) Huntington's disease and Huntington's disease-like syndromes: an overview. KCurr Opin Neurol 26(4):420-427. doi: 10.1097/WCO.0b013e3283632d90
5. Paterson AD1, Naimark DM, Huang J, Vachon C, Petronis A, King RA, Anderson VE, Sellers TA (1999) Genetic anticipation and breast cancer: a prospective follow-up study. Breast Cancer Res Treat 55(1):21-28.
6. Dagan E, Gershoni-Baruch R (2002) Anticipation in hereditary breast cancer. Clin Genet 62(2):147-150.
7. Kuchenbaecker KB, Hopper JL, Barnes DR, Phillips KA, Mooij TM, Roos-Blom MJ, Jervis S, van Leeuwen FE, Milne RL, Andrieu N, Goldgar DE, Terry MB, Rookus MA, Easton DF, Antoniou AC, BRCA1 and BRCA2 Cohort Consortium, McGuffog L, Evans DG, Barrowdale D, Frost D, Adlard J, Ong KR, Izatt L, Tischkowitz M, Eeles R, Davidson R, Hodgson S, Ellis S, Nogues C, Lasset C, Stoppa-Lyonnet D, Fricker JP, Faivre L, Berthet P, Hooning MJ, van der Kolk LE, Kets CM, Adank MA, John EM, Chung WK, Andrulis IL, Southey M, Daly MB, Buys SS, Osorio A, Engel C, Kast K, Schmutzler RK, Caldes T, Jakubowska A, Simard J, Friedlander ML, McLachlan SA, Machackova E, Foretova L, Tan YY, Singer CF, Olah E, Gerdes AM, Arver B, Olsson H (2017) Risks of Breast, Ovarian, and Contralateral Breast Cancer for BRCA1 and BRCA2 Mutation Carriers. JAMA 317(23):2402-2416. doi: 10.1001/jama.2017.7112.

8. Milne RL, Antoniou AC (2016) Modifiers of breast and ovarian cancer risks for BRCA1 and BRCA2 mutation carriers. Endocr Relat Cancer 23(10):T69-T84. doi: 10.1530/ERC-16-0277.
9. Kuchenbaecker KB, McGuffog L, Barrowdale D, Lee A, Soucy P, Dennis J, Domchek SM, Robson M, Spurdle AB, Ramus SJ, Mavaddat N, Terry MB, Neuhausen SL, Schmutzler RK, Simard J, Pharoah PDP, Offit K, Couch FJ, Chenevix-Trench G, Easton DF, Antoniou AC, Lush M, Hamann U, Southey M, John EM, Chung WK, Daly MB, Buys SS, Goldgar DE, Dorfling CM, van Rensburg EJ, Chun Ding Y, Ejlertsen B, Gerdes AM, Hansen TV, Slager S, Hallberg E, Benitez J, Osorio A, Cohen N, Lawler W, Weitzel JN, Peterlongo P, Pensotti V, Dolcetti R, Barile M, Bonanni B, Azzollini J, Manoukian S, Peissel B, Radice P, Savarese A, Papi L, Giannini G, Fostira F, Konstantopoulou I, Adlard J, Brewer C, Cook J, Davidson R, Eccles D, Eeles R, Ellis S, Frost D, Hodgson S, Izatt L, Lalloo F, Ong KR, Godwin AK, Arnold N, Dworniczak B, Engel C, Gehrig A, Hahnen E, Hauke J, Kast K, Meindl A, Niederacher D, Schmutzler RK, Varon-Mateeva R, Wang-Gohrke S, Wappenschmidt B, Barjhoux L, Collonge-Rame MA, Elan C, Golmard L; GEMO Study Collaborators; EMBRACE, Barouk-Simonet E, Lesueur F, Mazoyer S, Sokolowska J, Stoppa-Lyonnet D, Isaacs C, Claes KBM, Poppe B, de la Hoya M, Garcia-Barberan V, Aittomaki K, Nevanlinna H, Ausems MGEM, de Lange JL, Gomez Garcia EB, Hogervorst FB; HEBON, Kets CM, Meijers-Heijboer HE, Oosterwijk JC, Rookus MA, van Asperen CJ, van den Ouweland AMW, van Doorn HC, van Os TAM, Kwong A, Olah E, Diez O, Brunet J, Lazaro C, Teule A, Gronwald J, Jakubowska A, Kaczmarek K, Lubinski J, Sukiennicki G, Barkardottir RB, Chiquette J, Agata S, Montagna M, Teixeira MR, Kyung Park S; KConFab Investigators, Olswold C, Tischkowitz M, Foretova L, Gaddam P, Vijai J, Pfeiler G, Rappaport-Fuerhauser C, Singer CF, Tea MKM, Greene MH, Loud JT, Rennert G, Imyanitov EN, Hulick PJ, Hays JL, Piedmonte M, Rodriguez GC, Martyn J, Glendon G, Mulligan AM, Andrulis IL, Toland AE, Jenson UB, Kruse TA, Sokilde Pedersen I, Thomassen M, Caligo MA, Teo SH, Berger R, Friedman E, Laitman Y, Arver B, Borg A, Ehrancrona H, Rantala J, Olopade OI, Ganz PA, Nussbaum RL, Bradbury AR, Domchek SM, Nathanson KL, Arun BK, James P, Karlan BY, Lester J, Simard J, Pharoah PDP, Offit K, Couch FJ, Chenevix-Trench G, Easton DF, Antoniou AC (2017) Evaluation of Polygenic

Risk Scores for Breast and Ovarian Cancer Risk Prediction in BRCA1 and BRCA2 Mutation Carriers. J Natl Cancer Inst 109(7). doi: 10.1093/jnci/djw302.

10. Cui J, Luo J, Kim YC, Snyder C, Becirovic D, Downs B, Lynch H, Wang SM (2016) Differences of Variable Number Tandem Repeats in XRCC5 Promoter Are Associated with Increased or Decreased Risk of Breast Cancer in BRCA Gene Mutation Carriers. Front Oncol 6:92. doi: 10.3389/fonc.2016.00092. eCollection 2016.
11. Michailidou K, Lindström S, Dennis J, Beesley J, Hui S, Kar S, Lemaçon A, Soucy P, Glubb D, Rostamianfar A, Bolla MK, Wang Q, Tyrer J, Dicks E, Lee A, Wang Z, Allen J, Keeman R, Eilber U, French JD, Qing Chen X, Fachal L, McCue K, McCart Reed AE, Ghoussaini M, Carroll JS, Jiang X, Finucane H, Adams M, Adank MA, Ahsan H, Aittomäki K, Anton-Culver H, Antonenkova NN, Arndt V, Aronson KJ, Arun B, Auer PL, Bacot F, Barrdahl M, Baynes C, Beckmann MW, Behrens S, Benitez J, Bermisheva M, Bernstein L, Blomqvist C, Bogdanova NV, Bojesen SE, Bonanni B, Børresen-Dale AL, Brand JS, Brauch H, Brennan P, Brenner H, Brinton L, Broberg P, Brock IW, Broeks A, Brooks-Wilson A, Brucker SY, Brüning T, Burwinkel B, Butterbach K, Cai Q, Cai H, Caldés T, Canzian F, Carracedo A, Carter BD, Castelao JE, Chan TL, David Cheng TY, Seng Chia K, Choi JY, Christiansen H, Clarke CL; NBCS Collaborators, Collée M, Conroy DM, Cordina-Duverger E, Cornelissen S, Cox DG, Cox A, Cross SS, Cunningham JM, Czene K, Daly MB, Devilee P, Doheny KF, Dörk T, Dos-Santos-Silva I, Dumont M, Durcan L, Dwek M, Eccles DM, Ekici AB, Eliassen AH, Ellberg C, Elvira M, Engel C, Eriksson M, Fasching PA, Figueroa J, Flesch-Janys D, Fletcher O, Flyger H, Fritschi L, Gaborieau V, Gabrielson M, Gago-Dominguez M, Gao YT, Gapstur SM, García-Sáenz JA, Gaudet MM, Georgoulias V, Giles GG, Glendon G, Goldberg MS, Goldgar DE, González-Neira A, Grenaker Alnæs GI, Grip M, Gronwald J, Grundy A, Guénel P, Haeberle L, Hahnen E, Haiman CA, Håkansson N, Hamann U, Hamel N, Hankinson S, Harrington P, Hart SN, Hartikainen JM, Hartman M, Hein A, Heyworth J, Hicks B, Hillemanns P, Ho DN, Hollestelle A, Hooning MJ, Hoover RN, Hopper JL, Hou MF, Hsiung CN, Huang G, Humphreys K, Ishiguro J, Ito H, Iwasaki M, Iwata H, Jakubowska A, Janni W, John EM, Johnson N, Jones K, Jones M, Jukkola-Vuorinen A, Kaaks R, Kabisch M, Kaczmarek K, Kang D, Kasuga Y, Kerin MJ, Khan S, Khusnutdinova E, Kiiski JI, Kim SW, Knight JA, Kosma VM, Kristensen VN, Krüger U, Kwong A, Lambrechts D, Le Marchand L, Lee E, Lee MH, Lee JW, Neng Lee C,

Lejbkowicz F, Li J, Lilyquist J, Lindblom A, Lissowska J, Lo WY, Loibl S, Long J, Lophatananon A, Lubinski J, Luccarini C, Lux MP, Ma ESK, MacInnis RJ, Maishman T, Makalic E, Malone KE, Kostovska IM, Mannermaa A, Manoukian S, Manson JE, Margolin S, Mariapun S, Martinez ME, Matsuo K, Mavroudis D, McKay J, McLean C, Meijers-Heijboer H, Meindl A, Menéndez P, Menon U, Meyer J, Miao H, Miller N, Taib NAM, Muir K, Mulligan AM, Mulot C, Neuhausen SL, Nevanlinna H, Neven P, Nielsen SF, Noh DY, Nordestgaard BG, Norman A, Olopade OI, Olson JE, Olsson H, Olswold C, Orr N, Pankratz VS, Park SK, Park-Simon TW, Lloyd R, Perez JIA, Peterlongo P, Peto J, Phillips KA, Pinchev M, Plaseska-Karanfilska D, Prentice R, Presneau N, Prokofyeva D, Pugh E, Pylkäs K, Rack B, Radice P, Rahman N, Rennert G, Rennert HS, Rhenius V, Romero A, Romm J, Ruddy KJ, Rüdiger T, Rudolph A, Ruebner M, Rutgers EJT, Saloustros E, Sandler DP, Sangrajrang S, Sawyer EJ, Schmidt DF, Schmutzler RK, Schneeweiss A, Schoemaker MJ, Schumacher F, Schürmann P, Scott RJ, Scott C, Seal S, Seynaeve C, Shah M, Sharma P, Shen CY, Sheng G, Sherman ME, Shrubsole MJ, Shu XO, Smeets A, Sohn C, Southey MC, Spinelli JJ, Stegmaier C, Stewart-Brown S, Stone J, Stram DO, Surowy H, Swerdlow A, Tamimi R, Taylor JA, Tengström M, Teo SH, Beth Terry M, Tessier DC, Thanasitthichai S, Thöne K, Tollenaar RAEM, Tomlinson I, Tong L, Torres D, Truong T, Tseng CC, Tsugane S, Ulmer HU, Ursin G, Untch M, Vachon C, van Asperen CJ, Van Den Berg D, van den Ouweland AMW, van der Kolk L, van der Luijt RB, Vincent D, Vollenweider J, Waisfisz Q, Wang-Gohrke S, Weinberg CR, Wendt C, Whittemore AS, Wildiers H, Willett W, Winqvist R, Wolk A, Wu AH, Xia L, Yamaji T, Yang XR, Har Yip C, Yoo KY, Yu JC, Zheng W, Zheng Y, Zhu B, Ziogas A, Ziv E; ABCTB Investigators; ConFab/AOCS Investigators, Lakhani SR, Antoniou AC, Droit A, Andrulis IL, Amos CI, Couch FJ, Pharoah PDP, Chang-Claude J, Hall P, Hunter DJ, Milne RL, García-Closas M, Schmidt MK, Chanock SJ, Dunning AM, Edwards SL, Bader GD, Chenevix-Trench G, Simard J, Kraft P, Easton DF (2017) Association analysis identifies 65 new breast cancer risk loci. Nature doi: 10.1038/nature24284. [Epub ahead of print]

12. Milne RL, Kuchenbaecker KB, Michailidou K, Beesley J, Kar S, Lindström S, Hui S, Lemaçon A, Soucy P, Dennis J, Jiang X, Rostamianfar A, Finucane H, Bolla MK, McGuffog L, Wang Q, Aalfs CM; ABCTB Investigators, Adams M, Adlard J, Agata S, Ahmed S, Ahsan H, Aittomäki K, Al-Ejeh F, Allen J, Ambrosone CB, Amos CI, Andrulis

IL, Anton-Culver H, Antonenkova NN, Arndt V, Arnold N, Aronson KJ, Auber B, Auer PL, Ausems MGEM, Azzollini J, Bacot F, Balmaña J, Barile M, Barjhoux L, Barkardottir RB, Barrdahl M, Barnes D, Barrowdale D, Baynes C, Beckmann MW, Benitez J, Bermisheva M, Bernstein L, Bignon YJ, Blazer KR, Blok MJ, Blomqvist C, Blot W, Bobolis K, Boeckx B, Bogdanova NV, Bojesen A, Bojesen SE, Bonanni B, Børresen-Dale AL, Bozsik A, Bradbury AR, Brand JS, Brauch H, Brenner H, Bressac-de Paillerets B, Brewer C, Brinton L, Broberg P, Brooks-Wilson A, Brunet J, Brüning T, Burwinkel B, Buys SS, Byun J, Cai Q, Caldés T, Caligo MA, Campbell I, Canzian F, Caron O, Carracedo A, Carter BD, Castelao JE, Castera L, Caux-Moncoutier V, Chan SB, Chang-Claude J, Chanock SJ, Chen X, Cheng TD, Chiquette J, Christiansen H, Claes KBM, Clarke CL, Conner T, Conroy DM, Cook J, Cordina-Duverger E, Cornelissen S, Coupier I, Cox A, Cox DG, Cross SS, Cuk K, Cunningham JM, Czene K, Daly MB, Damiola F, Darabi H, Davidson R, De Leeneer K, Devilee P, Dicks E, Diez O, Ding YC, Ditsch N, Doheny KF, Domchek SM, Dorfling CM, Dörk T, Dos-Santos-Silva I, Dubois S, Dugué PA, Dumont M, Dunning AM, Durcan L, Dwek M, Dworniczak B, Eccles D, Eeles R, Ehrencrona H, Eilber U, Ejlertsen B, Ekici AB, Eliassen AH; EMBRACE, Engel C, Eriksson M, Fachal L, Faivre L, Fasching PA, Faust U, Figueroa J, Flesch-Janys D, Fletcher O, Flyger H, Foulkes WD, Friedman E, Fritschi L, Frost D, Gabrielson M, Gaddam P, Gammon MD, Ganz PA, Gapstur SM, Garber J, Garcia-Barberan V, García-Sáenz JA, Gaudet MM, Gauthier-Villars M, Gehrig A; GEMO Study Collaborators, Georgoulias V, Gerdes AM, Giles GG, Glendon G, Godwin AK, Goldberg MS, Goldgar DE, González-Neira A, Goodfellow P, Greene MH, Alnæs GIG, Grip M, Gronwald J, Grundy A, Gschwantler-Kaulich D, Guénel P, Guo Q, Haeberle L, Hahnen E, Haiman CA, Håkansson N, Hallberg E, Hamann U, Hamel N, Hankinson S, Hansen TVO, Harrington P, Hart SN, Hartikainen JM, Healey CS; HEBON, Hein A, Helbig S, Henderson A, Heyworth J, Hicks B, Hillemanns P, Hodgson S, Hogervorst FB, Hollestelle A, Hooning MJ, Hoover B, Hopper JL, Hu C, Huang G, Hulick PJ, Humphreys K, Hunter DJ, Imyanitov EN, Isaacs C, Iwasaki M, Izatt L, Jakubowska A, James P, Janavicius R, Janni W, Jensen UB, John EM, Johnson N, Jones K, Jones M, Jukkola-Vuorinen A, Kaaks R, Kabisch M, Kaczmarek K, Kang D, Kast K; kConFab/AOCS Investigators, Keeman R, Kerin MJ, Kets CM, Keupers M, Khan S, Khusnutdinova E, Kiiski JI, Kim SW, Knight JA, Konstantopoulou I, Kosma VM, Kristensen VN, Kruse TA, Kwong A, Lænkholm AV,

Laitman Y, Lalloo F, Lambrechts D, Landsman K, Lasset C, Lazaro C, Le Marchand L, Lecarpentier J, Lee A, Lee E, Lee JW, Lee MH, Lejbkowicz F, Lesueur F, Li J, Lilyquist J, Lincoln A, Lindblom A, Lissowska J, Lo WY, Loibl S, Long J, Loud JT, Lubinski J, Luccarini C, Lush M, MacInnis RJ, Maishman T, Makalic E, Kostovska IM, Malone KE, Manoukian S, Manson JE, Margolin S, Martens JWM, Martinez ME, Matsuo K, Mavroudis D, Mazoyer S, McLean C, Meijers-Heijboer H, Menéndez P, Meyer J, Miao H, Miller A, Miller N, Mitchell G, Montagna M, Muir K, Mulligan AM, Mulot C, Nadesan S, Nathanson KL; NBSC Collaborators, Neuhausen SL, Nevanlinna H, Nevelsteen I, Niederacher D, Nielsen SF, Nordestgaard BG, Norman A, Nussbaum RL, Olah E, Olopade OI, Olson JE, Olswold C, Ong KR, Oosterwijk JC, Orr N, Osorio A, Pankratz VS, Papi L, Park-Simon TW, Paulsson-Karlsson Y, Lloyd R, Pedersen IS, Peissel B, Peixoto A, Perez JIA, Peterlongo P, Peto J, Pfeiler G, Phelan CM, Pinchev M, Plaseska-Karanfilska D, Poppe B, Porteous ME, Prentice R, Presneau N, Prokofieva D, Pugh E, Pujana MA, Pylkäs K, Rack B, Radice P, Rahman N, Rantala J, Rappaport-Fuerhauser C, Rennert G, Rennert HS, Rhenius V, Rhiem K, Richardson A, Rodriguez GC, Romero A, Romm J, Rookus MA, Rudolph A, Ruediger T, Saloustros E, Sanders J, Sandler DP, Sangrajrang S, Sawyer EJ, Schmidt DF, Schoemaker MJ, Schumacher F, Schürmann P, Schwentner L, Scott C, Scott RJ, Seal S, Senter L, Seynaeve C, Shah M, Sharma P, Shen CY, Sheng X, Shimelis H, Shrubsole MJ, Shu XO, Side LE, Singer CF, Sohn C, Southey MC, Spinelli JJ, Spurdle AB, Stegmaier C, Stoppa-Lyonnet D, Sukiennicki G, Surowy H, Sutter C, Swerdlow A, Szabo CI, Tamimi RM, Tan YY, Taylor JA, Tejada MI, Tengström M, Teo SH, Terry MB, Tessier DC, Teulé A, Thöne K, Thull DL, Tibiletti MG, Tihomirova L, Tischkowitz M, Toland AE, Tollenaar RAEM, Tomlinson I, Tong L, Torres D, Tranchant M, Truong T, Tucker K, Tung N, Tyrer J, Ulmer HU, Vachon C, van Asperen CJ, Van Den Berg D, van den Ouweland AMW, van Rensburg EJ, Varesco L, Varon-Mateeva R, Vega A, Viel A, Vijai J, Vincent D, Vollenweider J, Walker L, Wang Z, Wang-Gohrke S, Wappenschmidt B, Weinberg CR, Weitzel JN, Wendt C, Wesseling J, Whittemore AS, Wijnen JT, Willett W, Winqvist R, Wolk A, Wu AH, Xia L, Yang XR, Yannoukakos D, Zaffaroni D, Zheng W, Zhu B, Ziogas A, Ziv E, Zorn KK, Gago-Dominguez M, Mannermaa A, Olsson H, Teixeira MR, Stone J, Offit K, Ottini L, Park SK, Thomassen M, Hall P, Meindl A, Schmutzler RK, Droit A, Bader GD, Pharoah PDP, Couch FJ, Easton DF, Kraft P, Chenevix-Trench G, García-Closas

M, Schmidt MK, Antoniou AC, Simard J (2017) Identification of ten variants associated with risk of estrogen-receptor-negative breast cancer. Nat Genet doi: 10.1038/ng.3785. [Epub ahead of print]

13. Trynka G, Westra HJ, Slowikowski K, Hu X, Xu H, Stranger BE, Klein RJ, Han B, Raychaudhuri S (2015) Disentangling the Effects of Colocalizing Genomic Annotations to Functionally Prioritize Non-coding Variants within Complex-Trait Loci. Am J Hum Genet 97(1):139-152. doi: 10.1016/j.ajhg.2015.05.016.
14. Bolton Ross JP, Grice DM, Bowden NA, Holliday EG, Avery-Kiejda KA, Scott RJ (2013) STaRRRT: a table of short tandem repeats in regulatory regions of the human genome. BMC Genomics 14:795.
15. Gerhardt J (2017) Epigenetic modifications in human fragile X pluripotent stem cells; Implications in fragile X syndrome modeling. Brain Res 1656:55-62. doi: 10.1016/j.brainres.2015.10.004.

Table 1. Matches between tandem repeats and significant SNPs for breast cancer among the general population, *BRCA1* and *BRCA2* mutation carriers

		General population				BRCA1 mutation carriers				BRCA2 mutation carriers			
Chr	**noTr**	**noSnp**	**Match**	**SnpIn**	**TrIn**	**noSnp**	**Match**	**SnpIn**	**TrIn**	**noSnp**	**Match**	**SnpIn**	**TrIn**
1	93640	16	69	14	69	4	17	4	17	4	13	4	13
2	93062	12	44	10	44	3	11	2	11	0	0	0	0
3	72794	10	32	10	32	0	0	0	0	2	11	2	11
4	72412	6	23	6	23	0	0	0	0	0	0	0	0
5	67737	16	77	14	77	3	17	3	17	4	23	4	23
6	65490	10	43	9	43	4	20	3	20	4	25	4	25
7	68593	8	28	6	28	1	2	1	2	0	0	0	0
8	57450	10	45	10	45	3	12	3	12	0	0	0	0
9	49378	8	28	8	26	1	2	1	2	0	0	0	0
10	55629	11	49	10	37	1	5	1	5	5	32	5	20
11	51145	6	41	6	41	2	13	2	13	3	16	3	16
12	55355	6	32	5	32	2	18	2	18	2	10	2	10
13	38032	2	4	2	4	1	2	1	2	0	0	0	0
14	35004	6	24	5	24	0	0	0	0	0	0	0	0
15	32212	1	3	1	3	0	0	0	0	0	0	0	0
16	41932	7	24	6	24	3	8	2	8	2	8	2	8
17	39419	5	20	5	20	0	0	0	0	2	6	2	6
18	29706	4	14	4	14	1	2	1	2	1	2	1	2
19	37498	7	44	7	44	1	9	1	9	1	5	1	5
20	27704	2	2	1	2	0	0	0	0	1	8	1	8
21	18086	1	0	0	0	0	0	0	0	1	0	0	0
22	19700	7	44	7	44	1	1	1	1	1	11	1	11
TOT	**1121978**	**161**	**690**	**146**	**676**	**31**	**139**	**28**	**139**	**33**	**170**	**32**	**158**

Legend- Chr = chromosome; noTr = number of TRs; noSnp = number of genome-wide significant SNPs; Match = number of matches between TRs and SNPs; SnpIn = number of 10kb-windows overlapping at least one TR; TrIn = number of TRs overlapping at least one 10kb-window.

Supplementary tables

General- Each table contains the following columns for the matches identified from the previously reported SNPs:

idSNP: SNP ID
Chr: Chromosome
staWin: Start position (Build 38) of the 10-kb window
posSNP: Position of the significant SNP
stoWin: Stop position of the 10-kb window
prGenes: Proposed candidate genes
idTR: Tandem repeat (TR) ID
staTR: Start position (Build 38) of the TR
stoTR: Stop position of the TR
lenPat: Size of the consensus pattern
noCop: Number of copies aligned with the consensus pattern
lenTR: Size of the TR
Pattern: Consensus pattern

Supplementary table 1A- Tandem repeats that co-localize to GWAS significant SNPs in the general population

Supplementary table 2A- Tandem repeats that co-localize to GWAS significant SNPs in *BRCA1* mutation carriers

Supplementary table 3A- Tandem repeats that co-localize to GWAS significant SNPs in *BRCA2* mutation carriers

Supplementary table 1B- Significant SNP's in BC GWAS in the general population

Supplementary table 2B- Significant SNP's in BC GWAS in *BRCA1* mutation carriers

Supplementary table 3B- Significant SNP's in BC GWAS in *BRCA2* mutation carriers